

---

# CONSTRUCTIONS WITH SHELL NOUNS IN ENGLISH: THEIR DUAL ROLE IN INFORMATION PACKAGING

## LES CONSTRUCCIONS ENCAPSULADORES EN ANGLÈS I LA DOBLE FUNCIO EN L'EMMAGATZEMATGE DE LA INFORMACIO

---

JARMILA TÁRNYIKOVÁ  
Palacký University Olomouc  
jarmila.tarnyikova@upol.cz

**Abstract:** The focus in this paper is on structural and functional properties of a marginal type of focus formulas in English, known as «*N-be-that-constructions*» (Schmid 2001), cf. *The problem is that he is jealous*. They are referred to here as *focus formulas with shell nouns* (FFSNs). The role of the nouns in these templates is to create conceptual shells (Schmid 1997) into which the propositional content is encapsulated, with both the parts being co-referentially linked. The BNC data have revealed two manifestations of FFSNs in discourse: they occurred either as fixed utterance-initial templates (*The thing is that killing got to be a habit.*), or as looser configurations, co-occurring with various discourse signposts (DSs), as in *Mm. well I mean there's the the other thing is you see...* My view, inspired by Schmid (2001) and outlined in Válková and Tárnýiková (2015), is that the FFSNs can have a dual role in discourse, operating either as focalizers or as discourse signposts participating in a number of pragmatic strategies associated with facework. The aim is to argue for the validity of this claim and give evidence of and reasoning for the focality loss cases.

**Key words:** focus formulas, discourse signposts, information packaging, shell nouns, grammaticalization, pragmaticalization.

**Resum:** Aquest treball se centra en les propietats funcionals i estructurals d'un tipus de fórmules focals en anglès: les construccions del tipus «*N-be-that*» (Schmid 2001), com en *The problem is that he is jealous*, a les quals denominem *fórmules amb encapsuladors* (FFSNs, segons la denominació en anglès). La funció dels noms en aquest patró consisteix a crear closques conceptuals (Schmid 1997) dins de les quals s'encapsula el contingut proposicional, una operació afavorida pel vincle coreferencial que hi ha entre les dues parts del patró. Les dades proporcionades pel BNC mostren dues manifestacions discursives de les FFSNs: ara apareixen com a patrons fixats d'inici

de seqüència (*The thing is that killing got to be a habit*), ara com a configuracions més laxes que concorren amb diverses fites textuais (*discourse signposts*, DSs), com en *Mm. well I mean there's the the other thing is you see...* La nostra perspectiva, inspirada en Schmid (2001) i esbossada en una recerca pilot (Válková & Tárnayiková 2015), és que les FFSNs poden tenir una funció discursiva dual, com a focalitzadors o com a atenuadors discursius que participen en una sèrie d'estratègies pragmàtiques associades amb la interacció cara a cara cordial i directa. L'objectiu d'aquest treball consisteix a comprovar empíricament aquesta dualitat.

**Paraules clau:** fórmules focals, fites textuais, emmagatzematge de la informació, encapsuladors, gramaticalització, pragmaticalització.

## 1. INTRODUCTION

### 1.1 THE IMPETUS TO STUDY FFSNS

There were two main reasons in my mind why to pay attention to shell nouns (SNs) and their role in discourse. First, as a functional linguist, I wanted to contribute to a scalar approach to the semantic saturation of nouns (Ns) in discourse and project SNs into an imaginary scale from semantically rich, i.e. full-content open class nouns in Schmid (1997: 5), to semantically unspecific nouns that can function as *tags*, e.g. in vague reference to notional categories (Channell 1994), as in *All about Cuckoos and Robins, and Things*, in which the tag *things* is used to activate the notional category of birds in our sample) — or *shells*, encapsulating the semantic content of the subsequent utterance, as in *The thing is [that people talk that way]*. Added to the series could also be vague non-numerical quantifiers in their figurative meaning (Tárnayiková 2010), such as *oceans of energy*, *bags of talent*, *mountains of evidence*, *a dash of courage*, and many more), not to speak about vague nouns used as euphemisms, such as Obama's frequent use of *thing* as a way of reference to invasion or war.

Put simply, the semantic saturation of nouns comes in degrees, and there are nouns and nouns, having their specific communicative values, ranging from semantically rich lexical items to tags or shells.<sup>1</sup>

1. Even the nouns functioning as conceptual shells, however, can reveal inner scalarity of semantic saturation (*thing* vs. *reason*), and inherently positive/negative evaluative connotations (*hope* vs. *problem*).

The second reason is a reaction to the strategy of some researchers to throw peripheral items or constructions under the carpet, or taxonomize them as *other possibilities* or *miscellany*. As prompted in an earlier contribution to the field (Válková and Tárnayiková 2015), I base my assumption on a dynamic approach to the centre-periphery hierarchy of language units, and advocate the idea that even marginal units or constructions can be endowed with specific communicative values associated with various facets of communicative maneuvering. Consequently, the status of FFSNs, compared to other ways of changing the information packaging in discourse (i.e. clefts, pseudo-clefts, extraposition, fronting, left- or right-dislocation), is peripheral but —as the immensity of corpus data has revealed— worth paying attention to, as the present study hopes to demonstrate.

#### 1.2 A TOP-DOWN PROCEDURE

Inspired by Halliday's architecture of language (Halliday & Webster 2009) and his three grammatically relevant metafunctions of the semantic system, i.e. ideational, interpersonal and textual, I approach FFSNs as overt language manifestations of Halliday's *textual* metafunction, whose primary role in discourse is to trigger the presupposition that the following part of information is focalized. As such, they fall within the category of *discourse signposts*, changing the information packaging. Emerging from the BNC data, however, are numerous FFSNs used as looser templates (going against the grain of expectations about the fixity of formulaic templates), sometimes almost lost in the surrounding clusters of discourse signposts.<sup>2</sup>

My strong inclination is to consider these looser configurations of FFSNs as overt language manifestations of the *interpersonal* rather than *textual* metafunction. The immensity of corpus data and emergent contextual factors optimizing either the focalizing or discourse signposting functions of FFSNs ought to help me find answers to the following research questions.

The semantic saturation within the cognitive space of SNs, however, has to be perceived as content with «structure-inherent semantic gaps» (Schmid 1997:8).

2. The term *discourse signposts*, borrowed from Válková (2012) is used here as an umbrella term for various discourse markers, hedges, mitigators, hesitation markers, such as *well, now, you see, I think, I mean, right*, and many more. Their more delicate taxonomy would be too specific for the purpose of the present study.

## I.3 RESEARCH QUESTIONS

(a) Why the template of the FFSNs, supposed to be structurally fixed (in order to serve the focalizing purpose in information packaging), appears in so many structural variables?

(b) What is the communicative effect of the frequent co-occurrence of FFSNs with discourse signposts (DSs), namely if the focus formulas and discourse signposts are interlaced into each other? (See sections 4.3 and 4.4.)

(c) Is my assumption about the dual role of FFSNs of general validity or is it just a potentiality partly turned into reality due to the semantic properties of SNs, their habituality, and/or the interplay of various factors of a given communicative situation?

## I.4 SHELL NOUNS AS PARTS OF A FOCALIZING CONSTRUCTION

Unlike the studies focusing on isolated tokens within the shell noun (SN) types, or studies foregrounding the cohesive role of SNs in discourse (Francis 1986; Flowerdew 2003), this paper approaches shell nouns as heads of a relatively fixed template (cf. the *N-be-that-construction* in Schmid 2000; 2001), whose focalising impact on the following proposition (the THAT-clause in our case) can be —due to various contextual factors— either strengthened or weakened. (For a detailed specification of the focalizing impact of SNs, see Schmid 2001, or Válková & Tárnayiková 2015.)

In order to collect more arguments by which to support the validity of a dynamic and context-sensitive approach to language data, my essential research pre-requisites include: the study of the structural properties of the focalizing construction (i.e. its constants and variables), the mapping of possible alternations within the SN space, the co-occurrence of SNs with utterance-initial discourse-regulative signposts, and the study of factors contributing to the fixity or flexibility of the FFSNs components.

Since SNs in the FFSNs operate as metadiscoursal containers linked to the encapsulated proposition, the semantic load of the encapsulated parts, realized in single propositions or in larger chunks of text (Schmid 1999: 116), also has a share in the overall focalizing effect of the FFSNs: with vague propositions, the focus formula can hardly fulfil its focalizing role, since there is in fact nothing to focus on (see 4.2-4.4). As Schmid (2001: 1529) put it, the focalizing presupposition triggered by the shell noun «can be a bluff». Put differently, not all the shells contain silvery pearls.

All these aspects will be taken into consideration in the application sections below, in which the BNC data will be used not only as supportive illustrative sam-

ples but before all as indicators of some tendencies creeping from the immensity of contextualized language use. First, however, the terminological apparatus has to be clarified and the status of the SNs delimited.

### 1.5 DIFFERENT APPROACHES AND UNSTABLE TERMINOLOGY

As with other language phenomena that can be approached from various perspectives, the labels used for what we discuss here as shell nouns, reflect various conceptions and should be approached as such. Below is a selective list of some of the approaches and terms, by which to illustrate the spectrum of views.<sup>3</sup>

For the purposes of this study, Schmid's (2000, 2001) approach will be prioritized together with his characteristics of SNs as quoted in 2.1. Selectively, I will take into consideration Stvan's (2007, 2014) results based on American English data (COCA), the findings of Delahunty (2011, 2012) and Kelzer (2013), and the tentative conclusions based on our pilot proposal (Válková & Tárníková 2015). Pragmatic considerations are mostly based on Leech (1983). The study is exclusively synchronic.

## 2. CHARACTERIZING SHELL NOUNS

Register for free at <https://www.scipedia.com> to download the version without the watermark

### 2.1 THE STATUS OF SHELL NOUNS

To a functional linguist, the most relevant and at the same time applicable characterization of SNs is in Schmid (2000). The author pinpoints the relatively

3. Francis speaks about *anaphoric nouns* (*A-nouns*) understood as signposts «by means of which he/she (reader) is periodically made aware of the writer's design and how the parts fit together in the development of the central theme» (1986: 2). In her revised versions (Francis 1994), the term *anaphoric nouns* gave way to *advance/retrospective labels*, obviously due to the need to cope with the cases in which the reference between the SNs and the encapsulated proposition is cataphoric; compare also her standpoint that «labels may function either cataphorically or anaphorically» (1994: 83). Ivanic prefers the term *carrier nouns*, stating that «they frequently carry a specific meaning within their context in addition to their dictionary meaning» (1991: 95). Hinkel's (2001, 2004) term *enumerative 'catch-all' nouns*, followed by the explanation that «[T]hey have specific identifiable referents in text, to which these nouns are connected» (2001: 129), reminds us of the role of SNs as one of the possible metaphors for containers (cf. also the container metaphors for vague quantity, such as *a bag of nerves*, which is close to Schmid's 2000, 2001 cognitively-based approach). Flowerdew's term *signalling nouns* and the explanation that «[A]ny abstract noun, the meaning of which can only be made specific by reference to its context» (2003: 2), are too general to delimit the class of shell nouns, though, implicitly, they are covered by the definition; compare also Schmid's view that abstract nouns are «suppliers of conceptual shells» (1997: 1).

open-ended nature of SNs, amplifying the potentiality of their status of conceptual shells and the scalarity of their nounhood (Schmid 2000: 4):

Shell nouns make up an open-ended functionally-defined class of abstract nouns that have, to varying degrees, the potential for being used as conceptual shells for complex, proposition-like pieces of information [...]. The property of shell-nounhood is a functional property.

In his later study, Schmid (2001: 1530) slightly modifies the open-ended status of shell nouns by stating that they «make up a limited set, though one with fuzzy edges».

Summing up, shell nouns are semantically rather unspecific conceptual shells, revealing various degrees of semantic saturation and operating within the FFSNs as heads used to trigger the presupposition that the more important piece of information (i.e. the shelled content) is to come. Hence my preference for the term *focus formulas with shell nouns* (FFSNs).

## 2.2 NARROWING THE SCOPE

In the application sections, the repertory of the relatively open-ended class of SNs will be restricted for the purpose of this study to include the following frequency-based samples of the BNC data: *thing* (6 667 results), *problem* (6 447), *point* (6 425), and *fact* (6 382). These four SNs are found to be the most frequent in the sample from the BYU-BNC data (accessed 2017-01-09). Emergent from the samples is the existence of the above mentioned potential spectrum of semantic saturation of SNs, ranging in our case from the semantically unspecific (vague) *thing* via inherently negative *problem*, inherently focal *point*, to objectivising *fact*. These four SNs will be approached from the perspective of their participation in both the relatively fixed and relatively flexible templates of FFSNs in the data.

## 3. STRUCTURAL PROPERTIES OF FFSNS

### 3.1 BASE STRUCTURE

Our starting point is the base structure of the FFSN, with its prototypical template

$$[((\text{Det}) + \text{SN} + \text{V}) + ((\text{THAT})\text{-clause})]$$

$$[\text{base space}] + [\text{focal space}],$$



in which Det stands for the determiner, SN is a shell noun, V is the linking verb BE, and THAT is the conjunction linking the *base space* [(Det) -SN -V] with the *focal space* [(THAT)-clause].

The presence or absence of the Det results in *full shells* or *bare shells*, with the latter restricted to a limited set of SNs; see Stvan (2014: 600-601) for her COCA-based list of bare shells said to visualise the process of grammaticalization in the inner structure of FFSNs, as in *Truth is, Trouble is, Thing is*. The reason why I include Det into the base form (though the bare shells are increasing in number) is that the bare form is a step to grammaticalization of the full form, similarly to juxtaposition being the result of THAT-deletion (*The thing is that he is lazy* > *The thing is he is lazy*).<sup>4</sup>

As for the encapsulated clause type, aside from the prototypical THAT-clause, the following clause types emerged from the BNC data: INF(itive)-clause (*The greatest thing in life is to keep your mind young* [BNC B2F\_commerce]), IF-clause (*The only thing I worry about is if I lose my mother at this time of life* [BNC FST\_W\_ac\_soc\_science]), HOW-clause (*Perhaps the best thing about Head Start 531 is how quickly you benefit* [BNCAYPWmisc]), or WH-clause (*But the thing is, what they do tend to do...* [BNC KCOSconv]). The focus in the present study is on the prototypical THAT-content clauses and their juxtaposed variants.

Rare but existing are the double FFSNs sequences, which might trigger the presupposition about a multiple (amplified) focalization—but, as the sample below prompts, this might be a red herring. If the juxtaposition is meant to move slow and vague (ex. (1) below), the initial double sequence of FFSNs (together with the discourse signpost *I mean*) only adds support to keeping the communicative line on in case of knowledge deficit, postponing bad news, or, for example, in waiting for a relevant thought to come. Such configurations, however, are always context-sensitive for their interpretation, cf.

- (1) The point is (unclear) (SP:KDAPSUNK) I mean the thing is were (unclear) some good conversation [BYU-BNC KDAS conv]

4. If we followed Cheshire's (1996: 384) claim that the function of *that* is «to coordinate the attention of speaker and addressee», then we would have to admit that in the juxtaposed clauses (with *that* deleted), the coordinating function is neutralized or suppressed. The lack of the connective, however, seems to be compensated for by the habituality of the juxtaposition and the resulting compactness between the SN and the focalized proposition.

## 3.2 TEMPLATEDNESS OF FFSNS (A SCALAR APPROACH)

In data processing, I activated both the *vertical*, paradigmatic axis of *alternation* (i.e. the choice within the spectrum of SNs used to initiate the formula), and the *horizontal*, syntagmatic axis of *co-occurrence* of the components within the FFSNs, together with the co-occurrence of the FFSNs with the surrounding discourse signposts, as in *Well the thing is I mean he plays for England* [BNC KD6 Sconv]). My first step was to consider the *inner configurations* of the FFSNs components and pay attention to corpus-emergent *constants* and *variables*. The constants correspond to the above-mentioned base structure [(Det)+SN+V + (THAT)-clause], as in

- (2) *The problem is that* the law is artificial [BYU-BNC A2P (692)]

The *base* structure is looked upon as a sequence with high degree of structural templatedness but with a whole spectrum of lexico-semantic alternations within the shell nouns.

The *variables* are potential modifications within each of the base components (see Table 1 below), changing the degree of templatedness in various ways: with Det, it is the presence or absence of the definite article or its alternation with possessive pronouns (e.g. *the/my problem is* but not *\*my fact is*; for some constraints on the definite article/possessive pronoun alternation, see Válková and Tárníková 2015: 70). The linking verb BE can occur in the present or past tense form, with the approximate ratio of 3:1 (based on 600 random BUY-BNC samples); THAT can be either explicit or implicitly deduced.

To illustrate the variables within the template, let us compare the following BYU-BNC samples in table 1, with the template in bold. The query was *Thing is* and *Thing was*, with *thing* intentionally chosen as the semantically most unspecific shell noun used by many authors as the prototype of shell nounhood (Delahunty 2011). The list below is selective, for more samples of structural variation see Válková and Tárníková (2015: 73).



	Data source	DSs	Det	Prem	shell	Postm	link	Con	Shelled content
1	BNC_CJRW_pop		The		thing		is	that	killing got to be a habit with Mario.
2	BNCADRW_bibliog.		The	important	thing	with this tour	is	that	it is personal
3	BNC_KD5S_conv				Thing		is	–	it's all right to tell people
4	BNC_KCVS_conv	Right, right (unclear)			thing		is	–	she's got her new pencil case.
5	BNC_JA2_S_interview	Yeah.	The	important	thing	for me to emphasize to you er	is	that	We don't make appointments.
6	BNC_JA9_S_meeting				Thing		was	–	we had lots of complaints last July

[BYU-BNC, accessed 2017-01-09]

Det= determiner, DS= discourse signpost, Con= connective, Prem= premodification, Post = post-modification

Table 1. Structural variety in the FFSN template with *thing*

With *fact*, however, the template is the strongest of all the four SNs, with only rare pre-modifications limited in our accessed data to signals of importance (*crucial*), or specification (*historical*). In 200 random samples of the focus formula with the SN *fact*, there was neither post-modification, nor interlacing with discourse signposts weakening the visualization of the template. Moreover, the base space of the formula is followed by a highly informative focal space (proposition) with a coherent content, so that the focalizing effect can be accomplished in the best possible way. The following table 2 will illustrate the situation.

	Data source	DSs	Det	Prem	shell	Postm	link	Con	Shelled content
1	BNC EF4 W_ commerce		The		fact		is	that	all the examples adduce to support the proposition that...
2	BNC_HHW hansard	(...) but	the		fact		is	that	it is the most treasured possession of many people
3	BNC_ISH S parliament		The		fact		is	that	the whole British business is now increasingly confident...
4	BNCC8R W ac_polit law edu		The	historical	fact		is	that	the settlement of 1688-1701 failed to settle everything;
5	BNC FBY W ac_polit edu		The	crucial	fact		is	that	if the school is over-subscribed, the parental wishes of some parents must...
6	BNC_ARK _W_fict_prose		The		fact		is	—	my husband is playing around with women.

[BYU-BNC, accessed 2017-01-09]

Table 2. Strong templatedness with the SN *fact*

Register for free at <https://www.scipedia.com> to download the version without the watermark

The preliminary conclusion we can draw is that we have to relativize our original hypothesis about the dual role of FFSNs in discourse by stating that not all the SNs are liable to the duality. *Point* almost mirrored the strong templatedness of *fact* in the accessed data, being a strong indicator of the focalizing function, while *problem* shared the duality of a strong and weak template prototypical of *thing*, hovering between a focaliser and a discourse signpost.

### 3.3 SHELL NOUNS OR SHELL NOUN PHRASES?

#### a. Premodification

The BNC samples reveal certain amount of flexibility in the pre-modification of SNs in FFSNs, which partly goes against templatedness but not necessarily against the focalizing effect of the FFSNs. As apparent from Table 3. below, visualising the frequency-based preferences in types of pre-modification of the SN *thing*, the pre-modifiers are short (longer pre-modifications of the type *the second and most important*,

*the only object-oriented, the biggest social*, are exceptional). Semantically, the pre-modifiers mostly concern enumeration (*first, second, other*), evaluation (*good, funny, worst*), or e.g. uniqueness (*only*); the superlatives (*most interesting, most important*) together with the intensifier *very* can put weight to the focalizing effect of the FFSNs.

Det + pre-modifier	Shell N	No of occur.	Det + pre-modifier	Shell N	No of occur.
The only	thing	1259	The second	thing	68
The first		868	The funny		66
The other		455	The good		33
The important		315	The second important		14
The most important		239	The very important		9

[BYU-BNC, accessed 2017-01-09]<sup>5</sup>

Table 3. Preferences in pre-modifying attributes of *thing*

#### b. Postmodification

The post-modification is more whimsical in nature, ranging from simple post-modifiers, as in

- (3) The good thing about the forest is that you can have a lot of people around... [BYU-BNC KIL Wnews script]

to long stretches, such as

- (4) The key thing in book selling, even more so than other forms of retailing is... [BYU-BNC 15 CBX Wcommerce],

5. With *problem, point, and fact*, there is also a preference for short pre-modifiers but their repertory and their frequency of occurrence significantly vary. The least pre-modified were *fact* (see table 2 above and additional pre-modifiers *important, most important, and second*), and *problem* (preferably pre-modified by *only, first, and second*). *Point* almost mapped the repertory of *thing* above, exclusive of the pre-modifier *funny*, (i.e. *\*funny point* or *\*funny fact*; cf. the constraints on modification mentioned above).

in which the long-distance separation of the linking verb from the SN(i.e. *thing---is*) decreases the templatedness of the FFSN and diminishes its focalizing effect. Similarly in (5) below

- (5) One very important thing that I had noticed since we landed in Normandy five days ago was that no one had any problem with insomnia [BYU BNC 6 A61 W biography]

### 3.4 STRUCTURAL PREFERENCES

Structural preferences in FFSNs can vary from shell noun to shell noun. Table 4 below surveys the frequency-based structural preferences with the shell nouns *thing* and *fact*, as emergent from the BYU-BNC data.

1.	Thing is	2,359	Fact is	777
2.	The thing is	671	The fact is	466
3.	Thing was	655	Fact is that	383
4.	Thing is that	375	The fact is that	276
5.	The thing was	91	Fact was	192
6.	The thing was that	65	The fact was	75
7.	Thing was that	8	The fact was that	50

[BYU-BNC accessed 2017-01-19]

(Different results emerged for *point* and *problem*.)

Table 4. Structural preferences with *thing* and *fact*

The high frequency of bare shells and deleted connectives in the FFSNs give support to the above mentioned process of *grammaticalization* within the FFSN template (Stvan 2014), as illustrated in 4.2a below.

## 4. TWO DIVERSE MANIFESTATIONS OF FFSNS IN DISCOURSE

### 4.1 LEFT-PERIPHERY DISTRIBUTION

FFSNs are prototypically located in the left periphery of the proposition (shaped as THAT-clause in our case, or a larger textual chunk). The left-periphery (i.e. the utterance initial position), however, is also the locus of various discourse signposts: discourse conjuncts (*but*, *and*...), words of agreement or disagreement (*yes*, *no*), and

various discourse markers regulating the smooth flow of communication. As a result, the co-occurrence of FFSNs and DSs —and the habitual status of both— can lead to various degrees of adaptability of FFSNs to the surrounding contextual milieu.

#### 4.2 GRADIENT OF THE FOCALIZING EFFECT

Emergent from the BNC data were two diverse structural manifestations of FFSNs, each having a scale of inner variants.

a. The FFSNs occurred as *relatively stable* utterance-initial templates (ex. (6)), with traces of grammaticalization visualized by the *reduction* of some constituents within the FFSNs (determiner, connective *that* or both), as in (7)-(8) below.

- (6) The problem is that the law is artificial. [BNC A2P (692)].
  - (7) The thing is  $\Delta$  this isn't a credit card is it? [BNC KCU Sconv]
  - (8)  $\Delta$  Thing is  $\Delta$  he don't care... [BNC KD7Sconv].
- (The delta symbol stands for the deleted component of the structure.)

The traces of grammaticalization in (7)-(8) are obvious, the question, however, is whether the process spreads to SNs as a type or to a limited number of tokens within the type. Stven (2014), basing her assumptions on the COCA data, argues for the latter alternative.

b. FFSNs emerged as looser configurations of the prototypical template, partly due to the process of *extension* (due to pre- or post-modifications of the SN within the template in Tables 1. and 2), as in *The important thing with this tour is that it is personal*. [BNC ADR W\_bibliogr]— and partly due to the co-occurrence of FFSNs with DSs, as in *Oh right! So, I mean (pause) the thing is, they are (pause) erm (pause) they are a minority group*. [BNC KDW S\_conv].

#### 4.3 SYMBIOSIS OF FFSNS WITH DISCOURSE SIGNPOSTS

The impact of the symbiosis of FFSNs and DSs on the focalizing or pragmatic role of the FFSNs depends on the *intensity* of the contact, the *length* of the DSs chains (clusters), and the way FFSNs and DSs are *mutually interwoven*. If the DSs sequentially precede the fixed template of the FFSN, the focalizing function of the formula is not

weakened, it is only shifted from the absolute beginning so that the focalizing effect is only *postponed* (ex. (9) below). In (10), however, the cluster of initial DSs is significantly long and the formula is followed by another DS, so that the FFSN, though shaped as a fixed template, is almost lost in the chain of various discourse signposts.

- (9) Yes, well the thing is I was afraid you might change your mind. [BNC HHA W fict prose].
- (10) Well, well, see how it goes, but erm I don't, the thing is you see they can teach you so much about it. [BNC KCY (112)].

#### 4.4 INTERLACING OF FFSNS WITH DISCOURSE SIGNPOSTS

In this section the focus is on those co-occurrences of FFSNs and DSs, in which the DSs are *interlaced* into the FFSNs, mostly inserted between the *base space* and the *focal space*, as identified in 3.1 above, or even spreading inside the proposition. Very often it is the case that all these possibilities occur in a complex interplay.

The following samples (11)-(14) are used to demonstrate the interlacing in which the visualization of the focus formula is weakened and the formula adapts to the chain of DSs to participate in various communicative maneuvers having to do with facework (i.e. the strategies not to lose one's own face and not to threaten the face of the other).

In (11), initiated by the prototypical contact marker *well*, we can trace the penetration of the chain of DMs in between the FFSN and the THAT-clause. The result is a long separation of the focus formula from the proposition and the FFSN is almost lost in the chain of DSs.

- (11) Well the thing is you see I think what from what I've gathered that... [BNC J8D(1996)]

In (12) the FFSN is preceded by two DSs and a pre-echoing Det (the). Following the FFSN is another DS so that the FFSN becomes part of a clusters of maneuvering discourse signposts after which the focalizing effect on the proposition (itself semantically unspecific) is suppressed.

- (12) Well the, I mean the thing is that I think that he is is... [BNC K71(366)]

In (13) the FFSN has a form of a fixed template but, as in (12), it is preceded by reiterated hesitation markers and followed by an incoherent torso of the proposition; there is hardly anything here that can be focalized, cf.

- (13) Er so er I mean the, the thing is that they, they, the... [BNC G4H (1403)]

Similarly in (14), the interplay of the initial sequence of discourse markers, the false start and the repetition of the SN determiner prompt hesitation, which, together with a weak information density of the propositional part of the utterance, neutralize the focalizing effect of the FFSN, though the formula again occurs in the format of a fixed template.

- (14) Mm. well I mean there's, the, the other thing is the differential around the, the county, cos there was a time when when... [BNC KLXS\_meeting]

The following examples (15)-(17), though sharing with the previous ones the co-occurrence of DSs with FFSNs, were primarily appended here to exemplify the situation in which, due to a vague proposition, the presupposition about focalization is a bluff.

In (15) the proposition is informatively so empty that the focus formula becomes a welcome way of prolonging the time of waiting for a thought to come or a welcome camouflage for knowledge deficit or the lack of coherent thinking.

- (15) Erm, the last thing is no, we, I don't want to stop the Chairman cos... [BNC D95\_meeting]

In (16) the initial chain of DMs, pauses, repetitions and false starts contribute to the overall hesitation effect in which the focal effect is not achieved by the FFSN (lost in the chain of discourse signpost) but rather by the end-focus position of the proposition.

- (16) Oh right! So, I mean (pause) the thing is, they are (pause) erm (pause) they are a minority group [BNC KDW S\_conv]

The complex interplay of false starts, discourse signposts, and incoherent proposition in (17) is illustrative of a contribution of the originally focalizing formula to the chain of devices by which to keep the communicative line while conveying almost nothing.<sup>6</sup>

6. In spoken interaction, the supra-segmental features, such as intonation contour or stress distribution, obviously play a relevant role in the decision-making processes of identifying the focalizing or discourse-signposting role of FFSNs, but their discussion exceeds the objectives of the present study.



- (17) Yeah and, and if you could, I mean the thing is, if you're, if er, if we didn't have each...  
[BYU-BNC KC7 S conv]

Summing up: among the *factors* influencing the process of pragmaticalization of FFSNs are: the in/flexibility of the structure of the FFSNs, the length of the clusters of DSs sequentially co-occurring with the formulas, the length of the clusters of DSs interlaced into FFSNs or even into the proposition, and last but not least, the semantic saturation of the proposition. If the propositional content of the THAT-clause is vague, and there is in fact nothing to draw attention to, the FFSN loses its focalising potential.

## 5. APPENDIX: ENGLISH-CZECH INTERFACE

In the following outline of the English-Czech interface, a selected number of cross-language differences will be discussed as a pre-requisite to my future plan, i.e. to focus on formal, functional and distributional properties of FFSNs in two typologically remote languages, English and Czech.

The preliminary comparison, based on the InterCorp section of the Czech National Corpus has confirmed that apart from the typologically predictable differences, there emerge from the data both quantitative and qualitative distinctions partly due to different conceptualization of the SNs domain, and partly due to different conventions in using SNs and consequent expectations of language users concerning the fixity and flexibility of the components within the FFSNs structure.

The following outline of less predictable differences should illustrate the case and the need to consider SNs within the scope of the whole focalizing formula and not only as isolated shells, since many variables within the shell NP are mutually interlinked with other elements of the focalizing construction (cf. the choice of V in particular NP-V collocations). The differences surveyed below were found crucial for the English-Czech comparison of *structural* properties, which does not mean that FFSNs have identical hierarchy within the system of focalizing constructions in the respective languages.<sup>7</sup>

7. In Czech, with a relatively free word-order, the focalizing (highlighting) effect can be prevalently achieved by word-order permutations, and the highlighting constructions (including focus formulas) tend to be perceived as stylistically marked options.

Our starting point will be the construction NP+BE+THAT-clause, as manifested in

- (18) E: The fact is that..., and its Czech counterpart  
Cz: Fakt je, že...

What is apparent at first sight is the absence of the definite article in the Czech version, which brings us to the first finding about the differences in the manifestation of the category of determination of SNs in the compared languages.

*Determination*, overtly signalled in English by definite/indefinite articles or a functional zero, is absent in Czech, since the category does not belong to the matrix of the grammatical categories prototypical of the Czech morphological class of nouns.

The following difference is in the fixity/flexibility of case marking.

*Case alternation in the SN form (nominative/instrumental)* — The form of the SN in Czech can alternate between the nominative and instrumental cases, cf.

- (19) Cz: Fakt<sub>[nom]</sub> je, že...[fact-is-that]  
Faktem<sub>[instr]</sub> je, že...[fact<sub>[instr]</sub>-is-that].  
E: The fact is that...

The alternation is perceived as neutral with some speakers, while some consider the nominative-alternation more colloquial and the instrumental-alternation more formal. The degree of formality, however, has to be considered in correlation with some other socio-pragmatic factors (e.g. degree of assertiveness, with instrumentals being perceived as less assertive). Moreover, the choice is, as mentioned below, partly governed by the collocability of shell Nouns with the following Verb (fakt<sub>nom</sub> je [fact is] vs. faktem<sub>instr</sub> zůstává [fact<sub>instr</sub> remains]).

*Singular/plural alternation in Czech SNs*, with singular forms dominating the possible choice, as in

- (20) Cz: Věc<sub>sg</sub> se má tak, že... /Věci<sub>pl</sub> se mají tak, že...  
[Thing itself has so that.../Things themselves have so that...]  
E: Thing is that...

The plural form of the SN, however, is not perceived by native speakers as signalling a sum of individual entities, i.e. many things, but rather as a quantum used to put weight to the seriousness, immensity or value of the proposition (similar e.g. to the plurals used to signal immensity rather than configuration of entities in vague non-numerical quantifiers (Channell, 1994) of the type oceans of energy, seas of blood, or bags of good luck).

*Fixity of the English V vs. a relative flexibility in the Czech V*

While there is a fixed use of BE in the English FFSNs, the Czech counterparts, on top of using the construction with BE (see above), reveal a limited, collocation-based flexibility in V, making use of such verbs as *remain*, *fall*, *have*, and *stand*, as in

- (21) Cz: Faktem<sub>[instr]</sub> je/zůstává  
 [fact<sub>[instr]</sub> is/remains]  
 E: The fact is
- (22) Cz: Padlo rozhodnutí  
 [fell- decision]  
 E: The decision was...
- (23) Cz: Problém se má tak, že...  
 [problem itself has so that]  
 E: The problem is that
- (24) Cz: Věc se má následovně (with semantically blocked that)  
 The thing itself has in the following way]  
 E: The thing is (as follows)
- (25) Cz: Problém/věc stojí tak, že...  
 [Problem/thing stands so that]  
 E: The problem/thing is...

The examples (21)-(25) above show that while in English the V belongs to the constants of the FFSNs template, in Czech it represents a variable, in which more specific verbs can alternate, creating collocability-based bonds with restricted types of SNs.

*Grammaticalization of the FFSN*

The process of grammaticalization, emergent from the BNC and COCA samples, as in *The thing is that he is lazy.* > *Thing is he is lazy.*, does not emerge from the Czech data.

## 6. CONCLUDING REMARKS

Let me summarize the points I have made. The BNC data (BYU-BNC) confirmed two manifestations of FFSNs in discourse and their dual role in information packaging. Both the manifestations, i.e. the fixed template and the looser configuration

interlaced with discourse signposts reveal two functions of FFSNs: their focalizing role in information packaging, and their pragmaticalized role in keeping the communicative line on and contributing to its smooth flow, while avoiding face threatening acts. The dual role of SNs in FFSNs construction lies in their dual contribution to communicative management: the strategies in information packaging in discourse and the strategies in facework. With this dual role of focalizers and discourse signposts, the FFSNs participate in the overt language manifestations of two metafunctions proposed by Halliday: textual (information packaging) and interpersonal (strategies monitoring the smooth flow of interaction).

The empirical part of my research, however, has weakened the general validity of my assumption about the dual role of FFSNs in discourse, since only some of the shell nouns are used in both fixed focalizing templates and flexible pragmatic enrichments symbiotically occurring with discourse signposts.

Further research, with a deeper insight into the *delicacy* within the proposed taxonomies will undoubtedly lead to revisions and consequent reformulations, to mode specification (spoken vs. written), variety specification (e.g. British vs. American English), genre specification, cross-language interface —and, hopefully, also to less tentative conclusions and more applicable results.

## BIBLIOGRAPHY

### Studies

- CHANNELL, J. (1994) *Vague language*, Oxford, Oxford University Press.
- CHESHIRE, J. (1996) «That jacksprat: An interactional perspective on English *that*», *Journal of Pragmatics*, 25, pp. 369-393.
- DELAHUNTY, G. (2011) «Contextually Determined Fixity and Flexibility in Thing Sentence Matrixes», in K. Kuiper (ed.), *Yearbook of Phraseology*, 2, Berlin, Mouton de Gruyter, pp. 109-136.
- (2012) «An Analysis of *The Thing is That* Sentences», *Pragmatics*, 22, 1, pp. 41-78.
- FLOWERDEW, J. (2003) «Signalling nouns in discourse», *English for Specific Purposes*, 22, pp. 329-346.
- FRANCIS, G. (1986) «Anaphoric nouns», *Discourse Analyses Monographs*, 11, Birmingham, University of Birmingham Printing Section.
- (1994) «Labelling discourse: an aspect of nominal group lexical cohesion», in M. Coulthard (ed.), *Advances in written text analysis*, London, Routledge, pp. 83-101.

- HALLIDAY, M. A. K. & J. J. WEBSTER (2009) *Continuum Companion to Systemic Functional Linguistics*, London / New York, Continuum International Publishing Group.
- HINKEL, E. (2001) «Matters of Cohesion in L2 Academic Texts», *Applied Language Learning*, 12 vol., 2, pp. 111-132.
- (2004) *Teaching Academic ESL Writing: Practical techniques in vocabulary and grammar*, New Jersey, Lawrence Erlbaum Associates.
- IVANIC R. (1991) «Nouns in search of a context: A study of nouns with both open- and closed system characteristics», *International Review of Applied Linguistics in Language Teaching*, 2, pp. 93-114.
- KELZER, E. (2013) «The X is (is) construction: An FDG account», in J. L. Mackenzie & H. Olbertz (ed.), *Coursebook in Functional Discourse Grammar*, Amsterdam/Philadelphia, John Benjamins.
- LEECH, G. N. (1983) *Principles of Pragmatics*, London, Longman.
- SCHMID, H. J. (1997) «Constant and ephemeral hypostatization: Thing, problem and other 'shellnouns'», in *Proceedings of the 16<sup>th</sup> International Congress of Linguistics (Paris, July 20-25)*. Elsevier, CD-ROM.
- (1999) «Cognitive effects of shell nouns», in Karen van Hoek, Andrej Kibrik & Leo Noordman (ed.), *Discourse studies in Cognitive Linguistics*, Amsterdam, Benjamins, pp. 111-132.
- (2000) «English Abstract Nouns as Conceptual Shells: From Corpus to Cognition», *Topics in English Linguistics*, Berlin / New York, Mouton de Gruyter.
- (2001) «Presupposition can be a bluff: How abstract nouns can be used as presupposition triggers», *Journal of Pragmatics*, 33, pp. 1529-1552.
- STVAN, L. S. (2007) «The Functional Range of Bare Singular Count Nouns in English», in E. Stark, E. Leill & A. Werner (ed.), *Nominal Determination: Typology, Context, Constraints, and Historical Emergence*, Amsterdam/Philadelphia, John Benjamins, pp. 171-187.
- (2014) «Truth Is, Sentence-Initial Shell Nouns Are Showing Up Bare», in L. Veselovská & M. Janebová, (ed.), *Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: «Language Use and Linguistic Structure»*. Olomouc Modern Language Series, 4 vol., Olomouc, Palacký University, pp. 591-606.
- TÁRNYIKOVÁ, J. (2010) «Bags of Talent, a Touch of Panic, and a Bit of Luck: the Case of Non-numerical Vague Quantifiers», *Linguistica Pragensia*, 20 vol., 2, pp. 71-85.
- VÁLKOVÁ, S. (2012) *Regulating Discourse: Compliments and Discourse Signposts (English-Czech Interface)*, Saarbrücken, LAP Lambert Academic Publishing.

VÁLKOVÁ, S. & J. TÁRNYIKOVÁ (2015) «Where focus formulas and discourse markers meet», *Discourse and Interaction*, 8 vol. 2, pp. 65-83.

Sources

*Czech National Corpus — InterCorp*. Prague, Institute of the Czech National Corpus. [Available online at <<http://www.korpus.cz>>.]

DAVIES, M. (2004-) *BYU-BNC*, (Based on the British National Corpus from Oxford University Press). [Available online at <<http://corpus.byu.edu/bnc/>>].

— (2008-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. [Available online at <<http://corpus.byu.edu/coca/>>].